National Climatic Data Center

DATA DOCUMENTATION

FOR

DATA SET 9669 (DSI-9669)

ESTIMATED MONTHLY USCRN STATION NORMALS OF TEMPERATURE, AND PRECIPITATION: 1971 - 2000

November 28, 2005

National Climatic Data Center
151 Patton Ave.
Asheville, NC 28801-5001 USA

:

Table of Contents

1. Abstract: The estimated normals presented in this publication include monthly maximum, minimum, and mean temperature and monthly total precipitation for the period 1971-2000 for stations of the U.S. Climate Reference Network (USCRN). In addition, statistics on the error of the estimated normal and its usefulness in operational climate monitoring is also provided. The normals were estimated using the monthly USCRN data and the adjusted monthly data from 4629 Cooperative (COOP) stations (Menne and Williams 2004). The methods used to estimate the normals and their error characteristics, and to evaluate the usefulness of estimated normals in climate monitoring were presented in great detail in Sun and Peterson (2005a, b), and were briefly described in Appendix A.

The USCRN started deploying stations in 2001. On a month by month basis only USCRN and COOP stations which had at least one month of data were included in the normals estimation of that particular month. As of December 2003, normals have been estimated for 42 USCRN stations, which are included in this dataset. This estimated normals product will be updated when more data are available and for more recently deployed USCRN stations.

Units used in this product are degrees F for temperature and inches for precipitation.

2. Element Names and Definitions: The data are archived in **two files** of fixed-length ASCII format.

File 1. ESTIMATED 1971-2000 MONTHLY NORMALS INVENTORY
(xxxx_1971-2000_NORM_CLIM??_MTH_STNMETA)
File 2. ESTIMATED 1971-2000 MONTHLY STATION NORMALS
(xxxxx_1971-2000_NORM_CLIM??_MTH_STNNORM)

File 1. ESTIMATED 1971-2000 MONTHLY NORMALS INVENTORY (xxxxx_1971-2000_NORM_CLIM??_MTH_STNMETA)

This file contains identification information about the stations for which 1971-2000 monthly normals were estimated.

ELEMENT	WIDTH	POSITION
STATION COOPERATIVE I.D. NUMBER (CD NUMBER)	6	001-006
CODE 1	1	007
CODE 2	1	008
STATION NAME	24	009-032
LATITUDE	5	033-037
BLANK	3	038-040
LONGITUDE	6	041-046
BLANK	3	047-049
ELEVATION	5	050-054
BLANK	2	055-056
CURRENT OBSERVATION TIME*	4	057-060
STATE ABBREVIATION	2	062-063

Code 1:
1 = Temperature only

:
:

2 = Precipitation only
8 = Temperature and Precipitation (No determination of snow)

Code 2:
0 = Station has at least one monthly value.
1 = Station has no monthly values.

*USCRN made measurements at each hour of a day from 0000 to 2300. The monthly USCRN data were calculated from data of midnight to midnight local time.

STATION COOPERATIVE I.D.: This 6-character station identifier (occasionally referred to as CD or Cooperative Number) is assigned by the National Climatic Data Center (POSITION 1-6). The first two digits refer to state code (Value 01 - 48, 50, 51, 66, 67 and 91). The next four digits refer to Cooperative Network Index Number (0001 - 9999) (Position 1-6).

DATA CODE 1: A one digit code (Position 7) in this column indicates whether a station is a temperature only (1), precipitation only (2), or temperature and precipitation (8).

DATA CODE 2: A one-digit code in this column indicates whether a station has at least one monthly value (0), or whether a station has no monthly values (1).

STATION NAME: An alpha, numeric or combination of both characters which indicate the station's name. Distance/direction addendums generally indicate number of miles and cardinal direction from a U.S. Post Office or centralized location associated with a place (e.g., NORTHPORT 2 W = 2 miles west of Northport Post Office or town center). A number of abbreviations are common, including: STN=Station, AP=Airport, INTL=International, NATL=National, RGNL=Regional, METRO=Metropolitan, OBSY=Observatory, UNIV=University, MTN=Mountain, ST PK=State Park, IS=Island, PLT=Plant, EXP=Experiment, REF=Refuge, AFB=Air Force Base, MCAS=Marine Corps Air Station, NAS=Naval Air Station (Position 9-32).

LATITUDE: In degrees and minutes (negative=S'ern hemisphere) (Pos. 33-37).

LONGITUDE: In degrees and minutes (negative=W'ern hemisphere) (Pos. 41-46).

ELEVATION: In whole feet (Pos. 50-54).

CURRENT OBSERVATION TIME: Local Standard Time at current observation time at which hour monthly USCRN values were calculated. (Pos. 57-60).

STATE ABBREVIATION: The 2-letter U.S. Postal Service abbreviation for states. Territories are assigned the following abbreviations: PR=Puerto Rico, VI=U.S. Virgin Islands, and PI=Pacific Islands (U.S. Pacific Trust Territories) (Pos. 62-63).
****

File 2. Estimated 1971-2000 MONTHLY STATION NORMALS (xxxxx_1971-2000_NORM_CLIM??_MTH_STNNORM)

:
:

This file contains the estimated 1971-2000 monthly station normals statistics.

ELEMENT		WIDTH	POSITION
STATION COOPERATIVE I.D. NUMBER (CD NUMBER)		6	001-006
ELEMENT CODE *		1	007
DATA CODE **		2	008-009
JANUARY DATA VALUE AND FLAG)))),		7	010-016
FEBRUARY DATA VALUE AND FLAG *		7	017-023
MARCH DATA VALUE AND FLAG *		7	024-030
APRIL DATA VALUE AND FLAG *		7	031-037
MAY DATA VALUE AND FLAG *	SEE	7	038-044
JUNE DATA VALUE AND FLAG /)	*** NOTE	7	045-051
JULY DATA VALUE AND FLAG *	BELOW	7	052-058
AUGUST DATA VALUE AND FLAG *		7	059-065
SEPTEMBER DATA VALUE AND FLAG *		7	066-072
OCTOBER DATA VALUE AND FLAG *		7	073-079
NOVEMBER DATA VALUE AND FLAG *		7	080-086
DECEMBER DATA VALUE AND FLAG *		7	087-093
ANNUAL DATA VALUE AND FLAG))))-		8	094-101

* ELEMENT CODE WHERE: 1 = MINIMUM TEMPERATURE
2 = MAXIMUM TEMPERATURE
3 = MEAN TEMPERATURE
4 = TOTAL PRECIPITATION

** DATA CODE WHERE:
03 = NUMBER OF MONTH/YEAR WITH USCRN DATA AVAILABLE
04 = ESTIMATED NORMAL
05 = 10 PERCENTILE ERROR OF ESTIMATED NORMAL
06 = 50 PERCENTILE ERROR OF ESTIMATED NORMAL
07 = 90 PERCENTILE ERROR OF ESTIMATED NORMAL
08 = PERCENTAGE OF MEDIAN ERROR DIVIDED BY TYPICAL MAGNITUDE OF YEAR-TO-YEAR VARIABILITY. IT IS USED TO MEASURE THE USEFULNESS OF ESTIMATED NORMAL IN OPERATIONAL CLIMATE MONITORING.

*** Note: DATA FIELD IS MADE UP OF VALUE (6 DIGITS) AND A FLAG BIT, EXCEPT ANNUAL VALUE. "-99999" INDICATES NO NORMAL AND ITS STATISTICS WERE CALCULATED.

FLAG, TO BE CONSISTENT WITH THE FORMAT OF MONTHLY STATION NORMALS PRODUCT (DSI-9641C), A BLANK (" ") IS USED.

TEMPERATURE UNITS = DEGREES AND TENTHS F (EXCEPT DEGREES AND THOUSANDTHS F FOR ERRORS)

PRECIPITATION UNITS = INCHES AND HUNDREDTHS (EXCEPT INCHES AND TEN-THOUSANDTHS FOR ERRORS)

STATION COOPERATIVE I.D.: This 6 character station identifier (occasionally referred to as CD NUMBER) is assigned by the National Climatic Data Center (POSITION 1-6). The first two digits refer to state code. (Value 01 - 48, 50, 51, 66, 67 and 91). The next four digits refer to cooperative Network Index Number (0001 - 9999) (Position 1-6).

:

:

ELEMENT CODE: Refers to Minimum/Maximum/Mean Temperature, Total Precipitation (Position 7).

DATA CODE: This code indicates what statistical parameter is listed on the record (Position 8-9).

MONTHLY DATA VALUE: 12 MONTHLY VALUES and one ANNUAL VALUE. Each month value consists of 6 digits/positions (Temperature-degrees and tenths F or Precipitation-inches and hundredths) plus 1 digit/position for a Flag Code. Position 10-16 January, Position 17-23 February,...., Position 80-86 November, Position 87-93 December and Position 94-101 ANNUAL VALUE and FLAG (7 digits/positions Annual value, 1 digit/position Flag Code) (Position 10-101).
****

---

3. **Start Date**: 19710101

4. **Stop Date**: 20001230

5. **Coverage**: The USA.

6. **How to Order Data**:

   Ask NCDC's Climate Services about the cost of obtaining this data set.
   Phone: 828-271-4800
   FAX: 828-271-4876
   E-mail: **NCDC.Orders@noaa.gov**

7. **Archiving Data Center**:

   National Climatic Data Center
   Federal Building
   151 Patton Avenue
   Asheville, NC 28801-5001
   Phone: (828) 271-4800.

8. **Technical Contact**:

   Dr. Bomin Sun
   National Climatic Data Center and STG Inc.
   Federal Building
   151 Patton Avenue
   Asheville, NC 28801-5001
   Phone: (828) 271-4865

   Dr. Michael R. Helfert
   National Climatic Data Center
   Federal Building
   151 Patton Avenue
   Asheville, NC 28801-5001
   Phone: (828) 271-4680

9. **Known Uncorrected Problems**: None.

:
:

10. Quality Statement: The monthly COOP dataset used in normals estimation has undergone extensive quality checks including range checks, estimation of missing data, evaluation of homogeneity, and adjustments of inhomogeneities. Quality of the U.S. Climate Reference Network data is considered quite good. All observations have received quality control.

11. Essential Companion Datasets: *Climatological Normals of the United States, DSI-9641C*
Monthly Station Normals of Temperature, Precipitation, and Heating and Cooling Degree Days, 1971-2000

DSI-3286, CLIMATE REFERENCE NETWORK

12. References:

Menne, M.J., and C.N. Williams, Jr., 2004: “1959 to 2003 serially complete adjusted monthly data documentation”, (for internal NOAA use only).

Sun, B., and T.C. Peterson, 2005a: “Estimating temperature normals for USCRN stations”, International Journal of Climatology (in press).

Sun, B., and T.C. Peterson, 2005b: “Estimating precipitation normals for USCRN stations”, Journal of Geophysical Research-Atmosphere (under review).

World Meteorological Organization, 1989: Calculation of Monthly and Annual 30-Year Standard Normals, WCDP-No. 10, WMO-TD/No. 341, Geneva: World Meteorological Organization.

:
:

Appendix A: Methodology of Normals Estimation

A climate normal is defined, by convention, as the arithmetic mean of a climatological element computed over three consecutive decades (WMO, 1989). USCRN was initialized in 2001, and as a result there are no calculated normals available. Because one of the goals of USCRN is to monitor weather and climate, estimated normals are needed. Methodologies used to estimate USCRN temperature and precipitation normals are presented in great detail in Sun and Peterson (2005a, b). A brief description is given in this note of the basic principle of the estimation and the evaluation on different estimation approaches. The evaluation acts as a role to find the best approach to estimating USCRN normals. Characteristics of errors generated from the best approach along with the usefulness of estimated normals in climate monitoring are also discussed in this appendix.

Principle of Normals Estimation

The monthly anomaly of a climatic parameter, such as temperature or precipitation, at a target location is generally similar to those neighboring stations. For temperature, this relationship can be expressed as a departure

$$(X - N)_{target} \approx (X-N)_{neigh} \quad (1.1)$$

where X are monthly values and N are normal values, respectively, and “target” and “neigh” stands for the target and neighboring stations, respectively.

For precipitation, the relationship can be described either as a departure (Eq. 1.1) or as a ratio

$$(X/N)_{target} \approx (X/N)_{neigh} \quad (1.2)$$

Solving these equations for N, the temperature normal at the target site can be estimated by

$$N_{target} \approx X_{target} - (X-N)_{neigh} \quad (2.1)$$

The precipitation normal can be estimated either by Eq. (2.1) or

$$N_{target} \approx X_{target} / (X/N)_{neigh} \quad (2.2)$$

Monthly Data

Two datasets were used in this study. The first is from the USCRN for which the normals were estimated. The monthly USCRN data were generated from daily data which were calculated from hourly values. The second is from the COOP network, which was used not only to estimate USCRN normals, but was also used to evaluate normals estimation approach and determine its error characteristics.

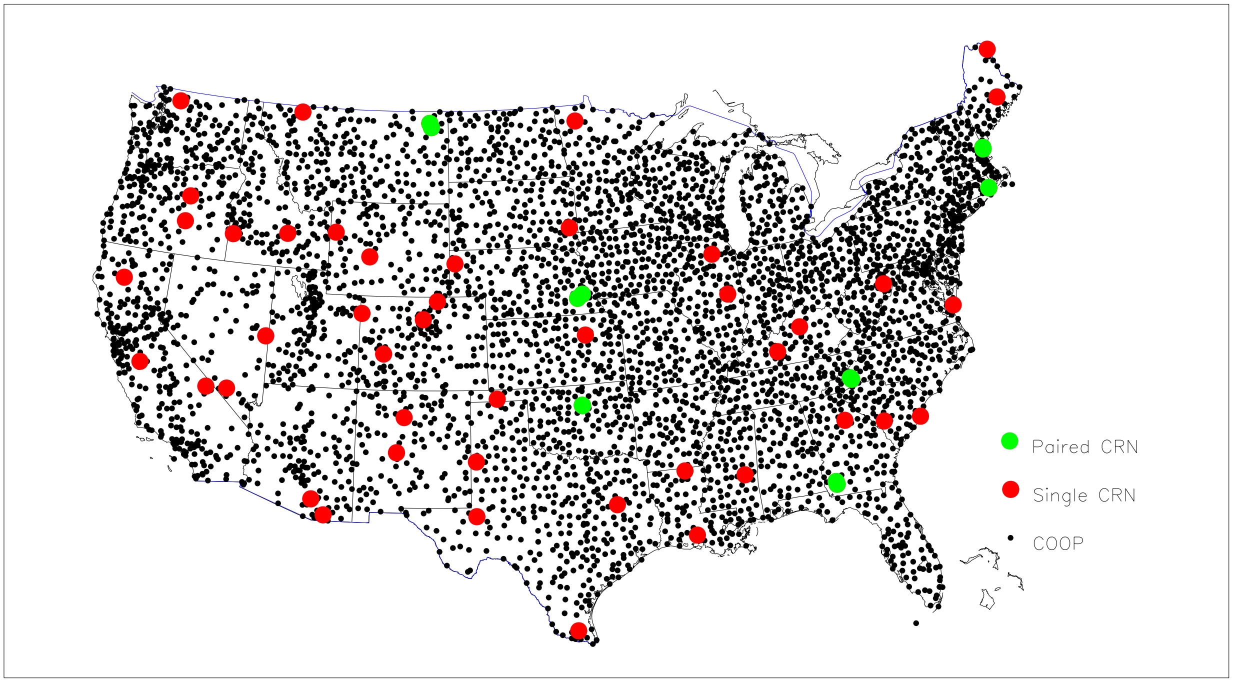


Figure 1. Data of 1971-2003 from 4629 COOP stations (Menne and Wililams, 2004) are used in normals estimation. As of July 2004, 60 USCRN stations were commissioned. The red ones stand for single USCRN stations and the green ones paired stations.

Sun and Peterson (2005a) evaluated several COOP datasets. All of them have actual normals (which are calculated from 1971-2000 data), but the homogeneity adjustment procedures applied to produce the datasets differ among them. It was found that errors of the estimated normals are sensitive to the homogeneity of COPOP datasets and the estimated normals generated from the 1959-2003 serially complete adjusted monthly dataset produced by M.J Menne and C.N. Williams at NCDC (Menne and Williams, 2004) are most accurate. Therefore, the Menne-Williams dataset, which contains the data from 4629 COOP stations (Fig. 1), were used to determine which estimation approach is best. The USCRN normals were then generated using the best approach from monthly USCRN data combined with monthly Menne-Williams data.

Evaluation on Estimation Approaches

Our goal is to have errors of the estimated normals as small as possible. Variations of several estimation approaches were evaluated, which provides the guidance to select the best approach. In this study, the error at a station is defined as the difference between the estimated and actual normal values.

Evaluation on Spatial Interpolation Scheme

In normals estimation, a spatial interpolation scheme is required to

:

:

interpolate the anomaly at a target station from anomalies of neighboring COOP stations. Fig. 2 shows errors of the normal estimated from three commonly used interpolation schemes: arithmetically averaging, inverse distance weighting, and Inverse Weighting of Square of temperature/precipitation Difference (IWSD) between the neighboring and target station. IWSD is a data-determined interpolation scheme that assigns more weight to the neighboring stations with observational values closer to the value at the target station.

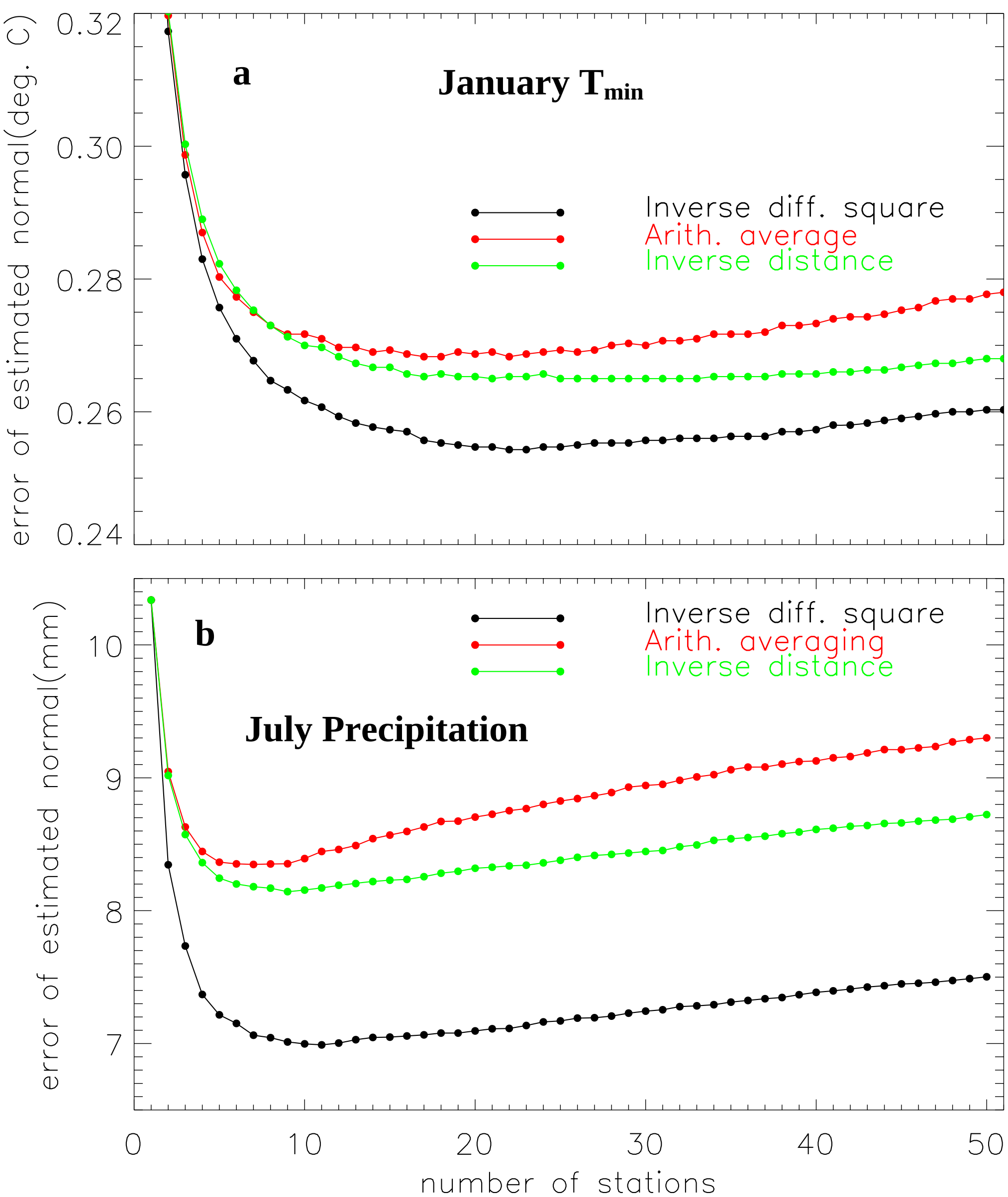


Figure 2. Sensitivity of error of estimated normals to spatial interpolation scheme. Normals were estimated by using 3-yr data. The errors shown are the overall median error values for the contiguous U.S. (a) January minimum temperature (T_{min}) and (b) July precipitation.

Comparison of the change of error with respect to the number of neighboring COOP stations shows a good agreement among the three schemes for temperature and precipitation, respectively (Fig. 2). Also note that the spatial pattern of error agrees well for the three schemes (not shown). IWSD, on the other hand, shows superiority over either of the other two schemes regardless of how many neighboring stations are used. For example, the error of estimated January minimum temperature (T_{min})normal associated with the use of 24 stations for IWSD is 0.25°C, about 5% smaller than the other two methods (Fig. 2a); the error of estimated July precipitation normal associated with the use of 11 stations is 7.06 mm, smaller than the other two methods by 16% (Fig. 2b). Similar results were obtained for other months for both temperature and precipitation.

IWSD is also better than other schemes we tested, including the Spatial Regression Test, an interpolation scheme developed for quality assurance purposes, and the SPHEREMAP algorithm (Sun and Peterson 2005a, b), an inverse distance weighting scheme which takes the directional distribution of stations into account.

Evaluation on Number of Neighboring Stations

"Neighboring stations" refers to nearby stations whose anomalies are used to estimate the normals at a target station. Errors in T_{min} and T_{max} of both January and July (Fig. 3a) show similar variations with respect to the number of neighboring stations used: they decrease rapidly with the increase of neighboring stations from one to around five, continue to decrease but more gradually, then reach a minimum value and afterwards increase slowly. The number of neighboring stations corresponding to the minimum error, the so called optimal number, varies slightly with month and T_{min} and T_{max}, but is around 24.

Similar changes are noticed in the errors of estimated precipitation normals (Fig. 3b), but the optimal number is around 11.

The errors in Fig. 3 represent the characteristics of median errors for the contiguous U.S. One might expect that the optimal number should vary with geographic regions or climate regimes. After conducting calculations for all of the 4629 COOP stations, we noticed that there are no strong regional patterns of the optimal number. In fact, across the country the optimal number varies strongly with location. In this study, data from 24 and 11 neighboring stations described above were selected to be used in normals estimation at all locations and in all months for temperature and precipitation, respectively.

:
:

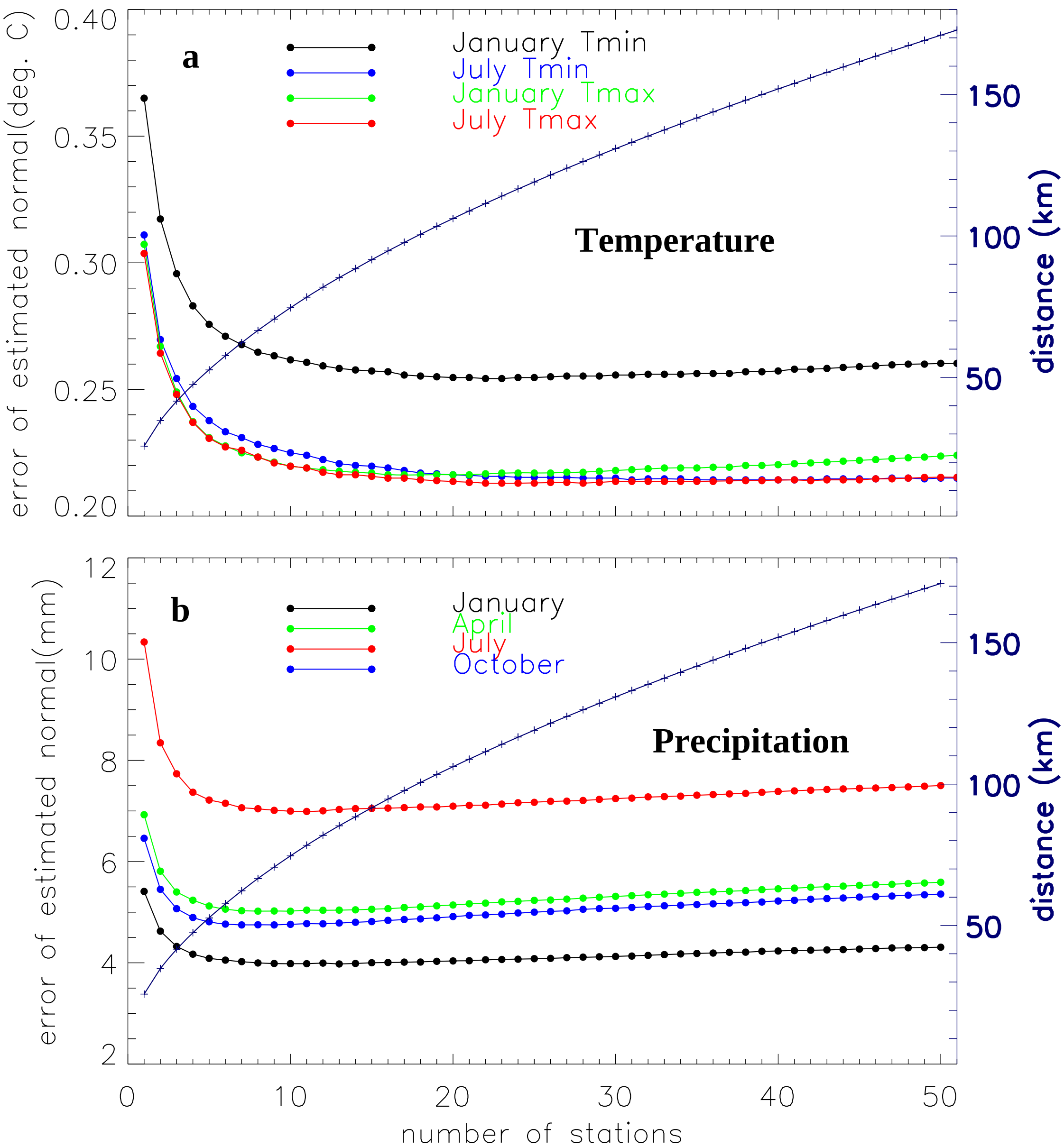


Figure 3. Sensitivity of error of estimated normals to the number of neighboring stations used. A weighting scheme involving the inverse of square difference in observational value (IWSD) between the neighboring and target station was used in the estimation. The normals were estimated by using 3-yr data. The errors shown are the overall median errors for the contiguous U.S. The "+" line on the right-y axis indicates the distance within which the neighboring COOP stations are located. (a) temperature and (b) precipitation.

Evaluation on Departure versus Ratio Method

As discussed in "Principle of Normals Estimation", a precipitation normal can be estimated by either using the departure (Eq. 2.1) or the ratio (eq. 2.2)

:

:

method. The normal, however, can not be estimated using the ratio method if the monthly precipitation total at the neighboring COOP station is zero, which is the case over some areas of the western U.S. The dataset used in this comparison therefore includes only those stations which have non-zero monthly precipitation values.

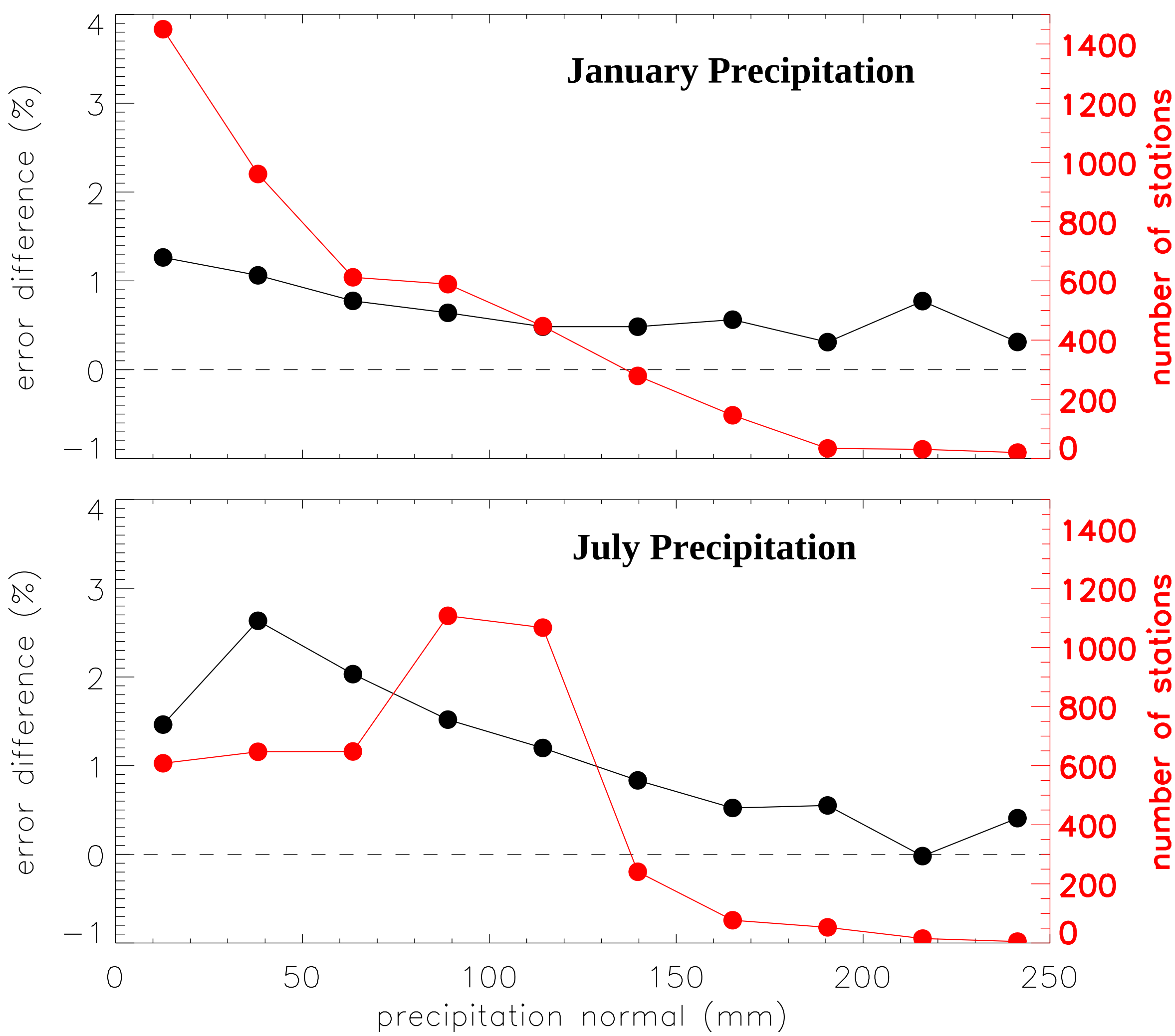


Figure 4. Difference in error of estimated precipitation normal between the ratio and the departure method. The x-axis represents the precipitation regimes categorized by normal values. The error values are the median ones determined from values of all of the stations shown on the right-y axis. The errors are estimated using 3-yr data and in percent of the normal values.

Fig. 4 indicates that errors estimated using the ratio method are greater than those using the departure method for all different precipitation regimes across the contiguous U.S. These comparisons lead us to decide that the departure method should be used in our normals estimation, although the ratio method has been widely used in similar applications. Please note errors shown in Figs. 2b and 3b were estimated using the departure method.

:

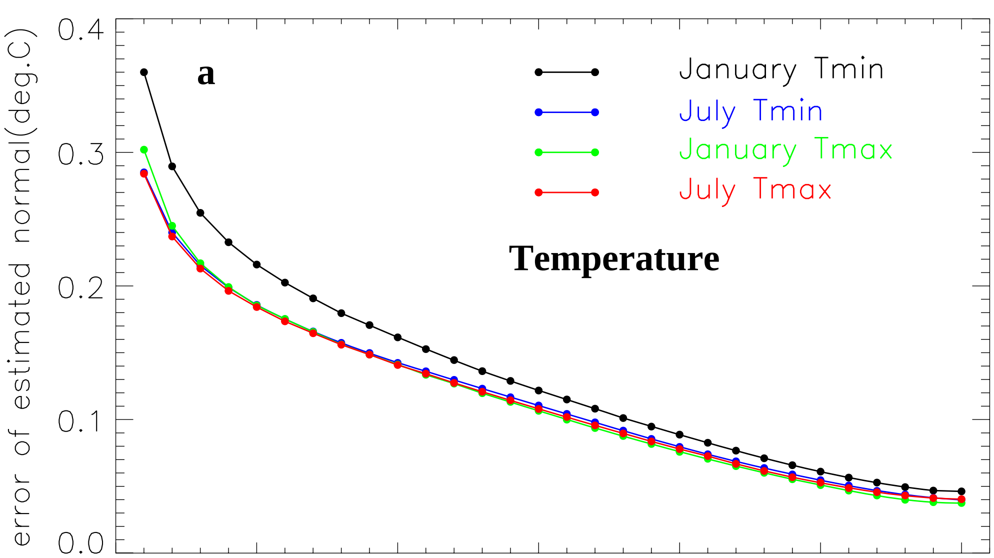
We are aware that a negative precipitation normal value can be generated from the departure method if the precipitation anomaly interpolated from the neighboring COOP stations is greater than the one at the target station. Our survey shows that stations with negative estimation account only for a very small portion of total stations. For example, this occurred at 0.2% of total stations when 3-yr data used. For those limited cases the normals were estimated using the ratio method.

To summarize, the best approach for temperature normals estimation includes the use of data from 24 neighboring stations and an interpolation scheme of ISWD, and the best approach for precipitation normals estimation includes the use of data from 11 neighboring stations and the interpolation scheme of ISWD. These approaches were used to generate error characteristics of estimated USCRN normals.

Results

Error Characteristics of Estimated Normals

As expected, for estimated normals of both temperature and precipitation (Fig. 5), error values decrease with the increasing number of years of data used; and the errors, however, are reduced faster when the starting number of years of data used is only a few. Seasonal variability is also noticed in error values. For temperature (Fig. 5a), errors in January T_{min} are greater than errors in July and errors in T_{max}; for precipitation (Fig. 5b), the greatest errors are found in July. Due to the lack of mixing, values of T_{min} in cold seasons are strongly affected by local siting features. The warm season precipitation is primarily dominated by small-scale convective processes. The greater errors in January T_{min} and July precipitation probably arise from their less spatial representativeness.



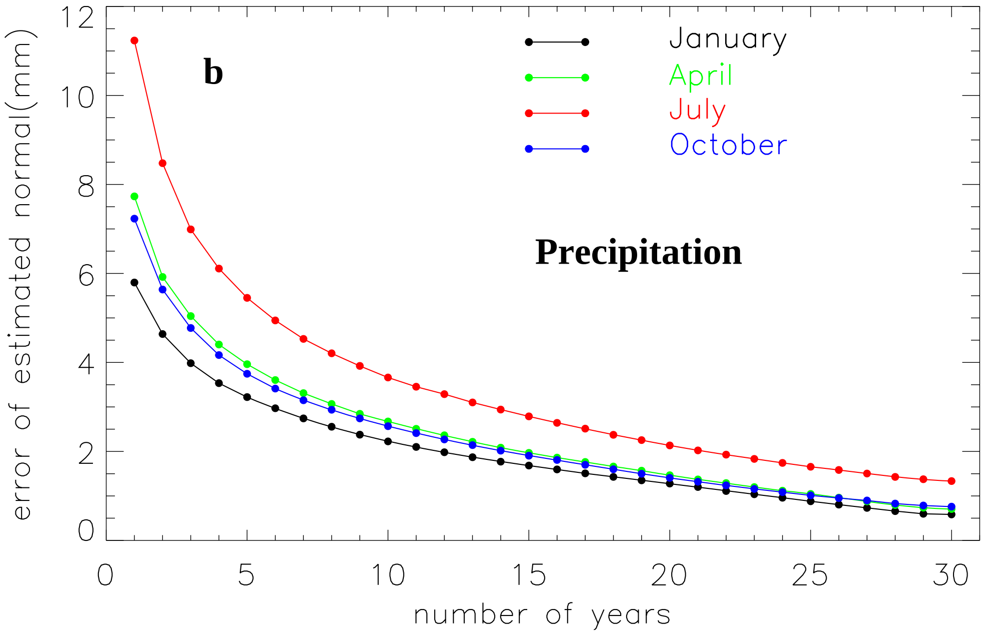


Figure 5. Change of the error with the number of years of data used. Temperature normals were estimated using data from 24 neighboring stations and the interpolation scheme of IWSD, and precipitation normals were estimated using data from 11 neighboring stations and the scheme of IWSD. The errors shown are the overall median errors for the contiguous U.S. (a) temperature and (b) precipitation.

Besides the median values, other percentile error values show similar changes with the number of years of data used. Examples are shown in Fig. 6. The 90% confidence intervals, which are the differences between the 95th and 5th percentile values, also decrease with more years of data used. For example, the 90th confidence interval for January T_{min} is 1.39°C and 0.94°C when using 1-yr and 3-yr data respectively, and is reduced to 0.78°C and 0.58°C when using 5-yr and 10-yr data, respectively; for July precipitation the 90th confidence interval is 52 mm and 30 mm when using 1-yr and 3-yr data respectively, and becomes 22 mm and 15 mm when using 5-yr and 10-yr data, respectively.

:
:

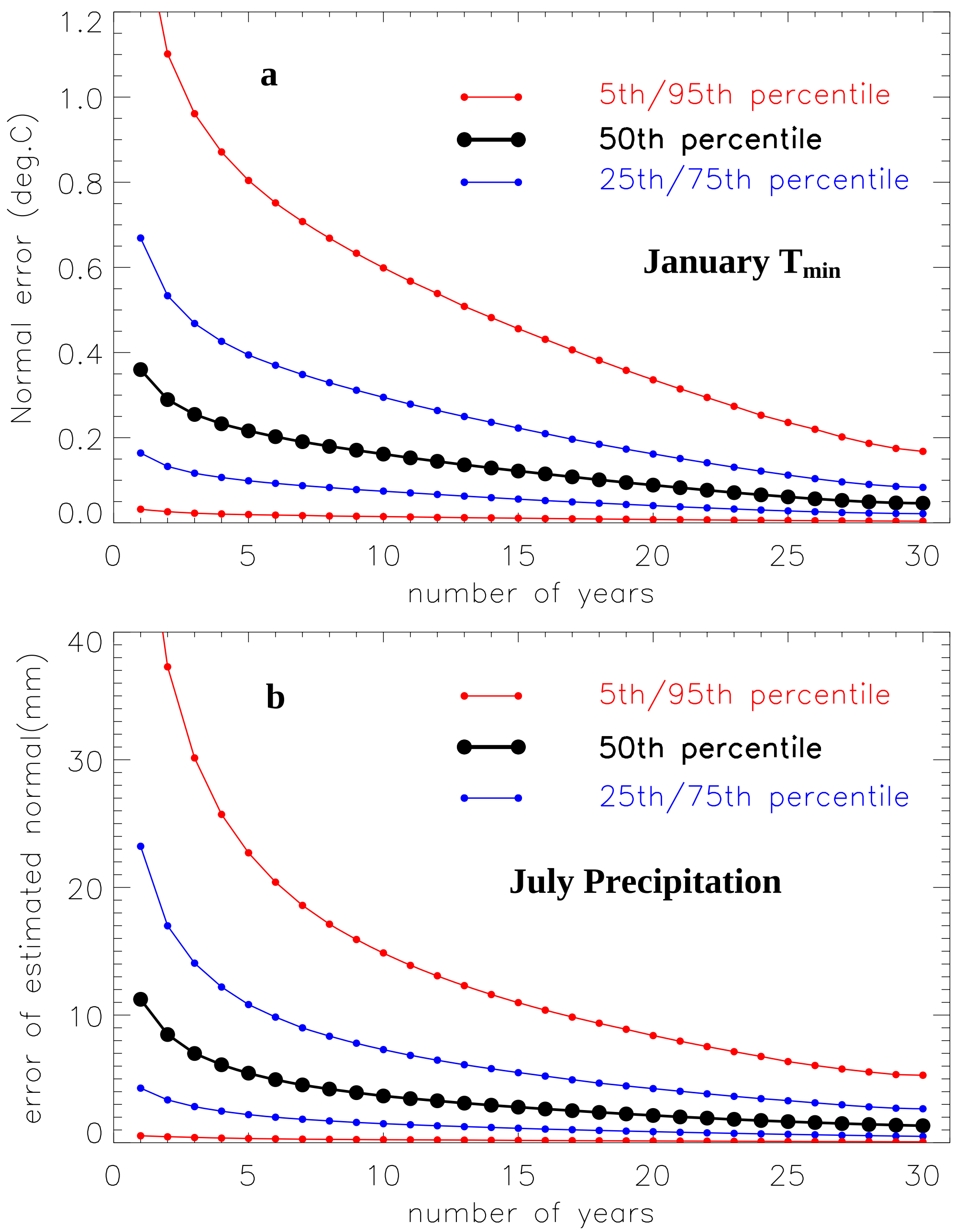


Figure 6. Changes of percentile error values with the number of years of data used. Normals were estimated using the same approaches as for Fig. 5. (a) January T_{min} and (b) July precipitation.

Errors of estimated normals vary not only with the number of years of data used but with location as well. Fig. 7 is examples of the spatial distribution of errors estimated using 3-yr data. As demonstrated in Sun and Peterson (2005a,b), the spatial pattern of error varies with month for both temperature and precipitation. In general, for temperature normals estimation, errors in the western U.S. are greater than in the eastern part; for precipitation normals estimation, smaller error values are found over dry areas and greater errors over wet areas. The magnitude of errors is reduced dramatically across the contiguous U.S. when more years of data are used.

Circles on the maps (Fig. 7) represent the location of USCRN stations deployed by July 2004. Please be noted that errors for estimated USCRN normals are assumed to be the same as those of neighboring stations and are therefore assigned from error values calculated using COOP data alone.

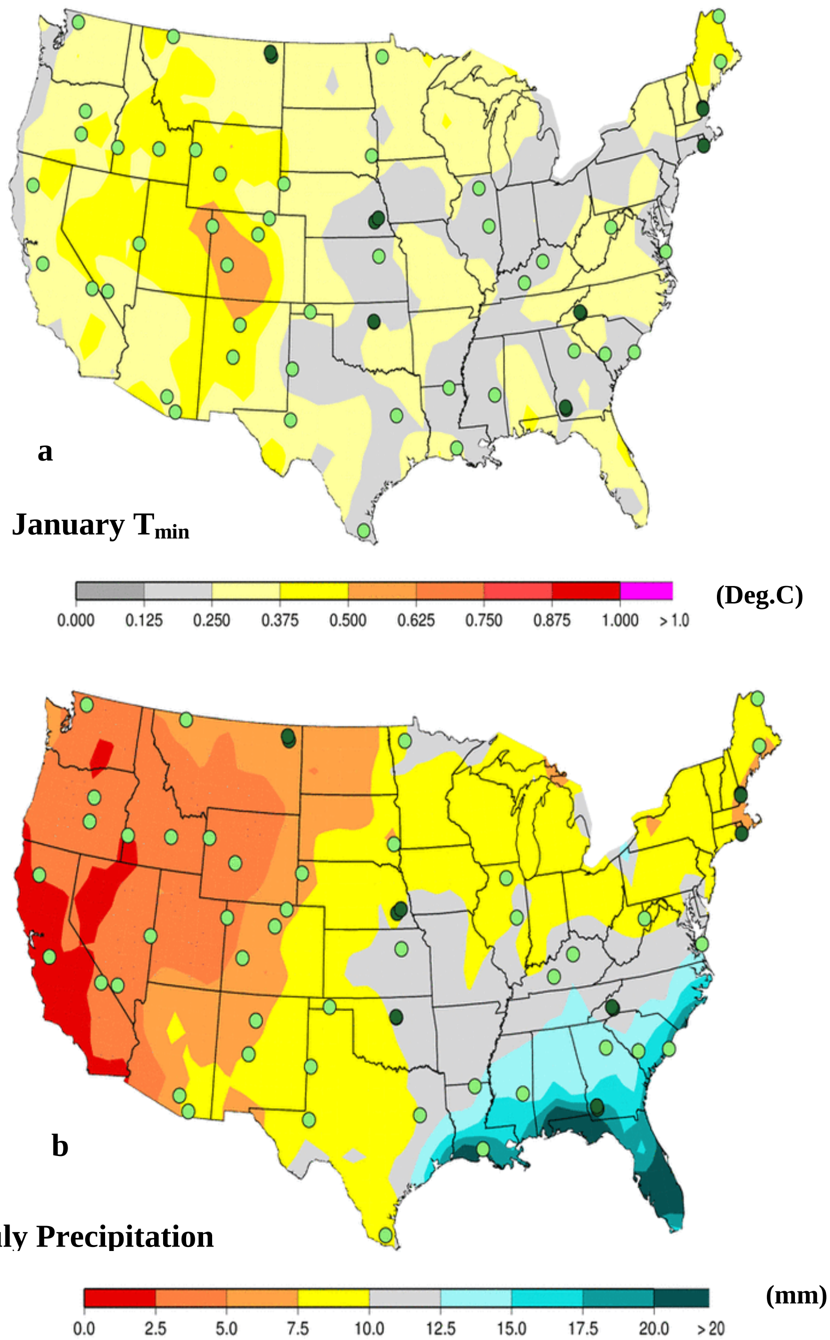


Figure 7. Errors of estimated normals for (a) January T_{min} and (b) July
:

precipitation. Normals were estimated from 3-yr data using the same approaches as the ones for Fig. 5.

Usefulness of Estimated Normals in Climate Monitoring

To be useful in climate monitoring activities, an estimated normal should have the error value smaller than the magnitude of the typical anomaly being monitored. The ratio of "error-to-anomaly" is employed to assess the usefulness of our estimated normals. Here the "anomaly" is the median value derived from the COOP anomaly values of 1971-2003. It is used to measure the typical magnitude of year-to-year variability.

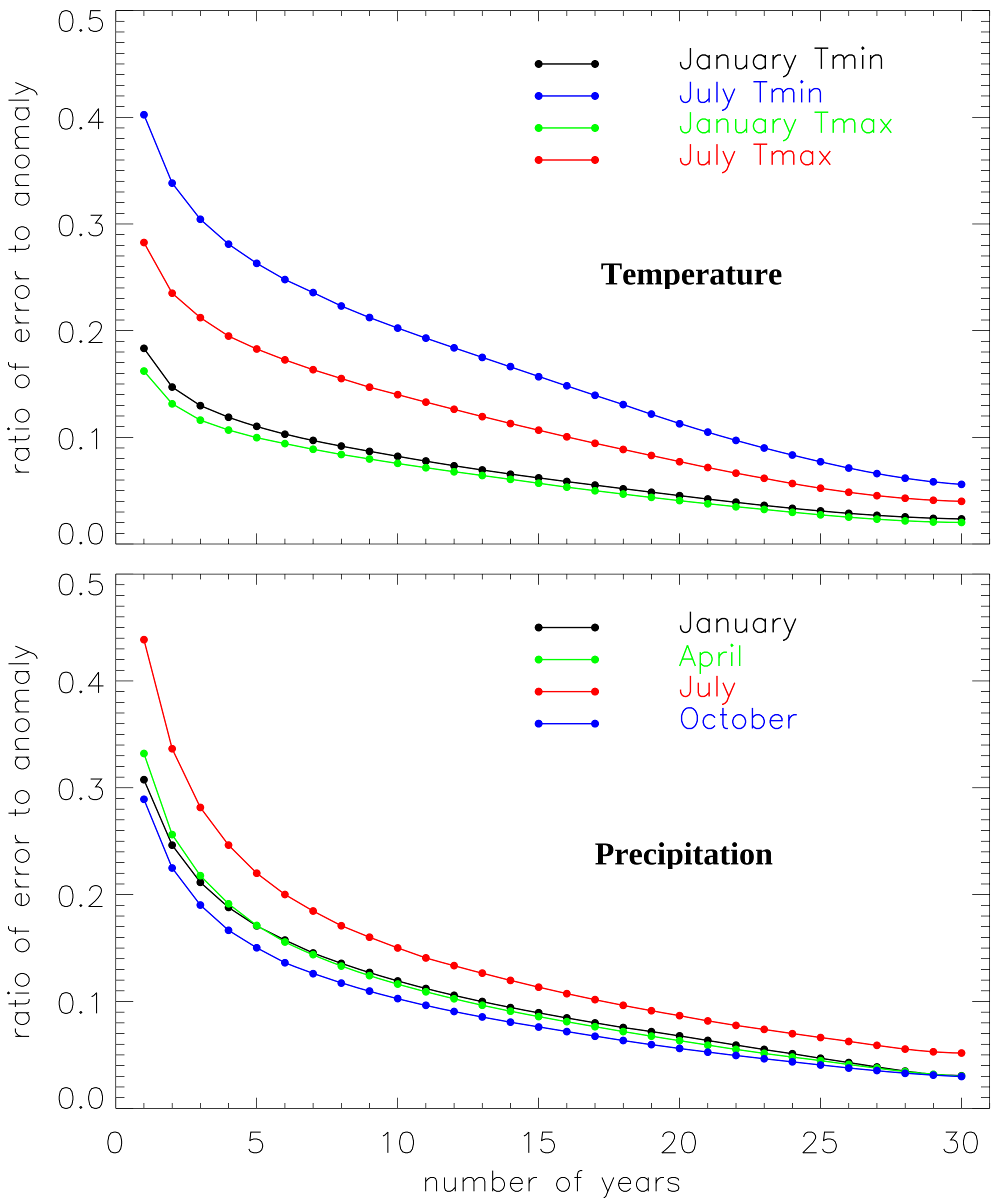


Figure 8. Ratios of error of estimated normal to typical anomaly value. Normals associated with the error were estimated using the same approach as

:

:

the one for Figure 5. The typical anomaly value measures the magnitude of year-to-year variability and is calculated from COOP data of 1971-2003.

:

:

Fig. 8 revealed that the error-to-anomaly ratio values decrease with more years of data used for both temperature and precipitation in the way their errors do. As expected, the ratio values for precipitation are generally greater than those for temperature particularly when a few years of data are available for normals estimation. Interestingly, the ratios are quite small if only a few years of data used. For example, if the normals are estimated using 3-yr data, the ratio is 0.13 for January T_{min} and 0.21 for July T_{max}; and the ratio is 0.21 for January precipitation and 0.28 for July precipitation. These numbers suggest the estimated USCRN normals can be useful in operational climate monitoring.

Summary

Normals estimation was based on the fact that monthly anomalies at any target location are similar to those of neighboring stations. Normals have been generated for USCRN stations by using data from this network and 4629 COOP stations. Based upon the evaluation on several variations of estimation approaches, the best approach used in USCRN normals estimation for temperature was one that included the use of monthly departure data from 24 neighboring COOP stations (within 117 km of a USCRN station), and the weighting scheme of ISWD; and the best approach used in USCRN precipitation normals estimation included the use of 11 neighboring stations (within 78 km of a USCRN station) and the weighting scheme of ISWD. When the estimated normals are based on a few years of data (e.g., 3-yr of data), the errors in estimated temperature normals are 12-30% of and the errors in estimated precipitation normals are about 10-14% of the typical magnitude of year-to-year variability. This suggests that these estimated USCRN normals can be useful in climate monitoring and assessment.

Temperature and precipitation normals have been estimated on a month basis for USCRN stations which have at least one month of data for that month. The annual normals have been estimated for USCRN stations which have at least one full year (twelve calendar months beginning January, ending December) of data.

:

:

21:

:
: